THE REGULARITY LEMMA WITH BOUNDED VC DIMENSION

HENRY TOWSNER

1. Introduction

In this note we give a proof of Szemerédi's celebrated regularity lemma [10] in the special case where the graph has bounded VC dimension. In the general case it is known that the tower exponential bounds given by Szemerédi's proof are essentially optimal [2], but in this special case we obtain doubly exponential bounds. After placing this online, we learned that a stronger result had already been obtained by Lovász and Szegedy [6], giving polynomial bounds (and a stronger condition on the partition). We leave this note online since the proof may still be of interest.

The results here are reminiscent of recent results by Malliaris and Shelah [7] obtaining improved bounds for the regularity lemma under various model theoretic assumptions (including bounded VC dimension, under its model theoretic name, NIP [5]). However their results focus on eliminating or controlling irregular pairs, while the result here keeps the irregular pairs but requires fewer components in the partition.

2. ϵ-Approximations

Throughout this note, we will be concerned with large finite sets X and Y. We will use measure-theoretic notation for the normalized counting measure: when $X' \subseteq X$, $\mu(X') = \frac{|X'|}{|X|}$, when $Y' \subseteq Y$, $\mu(Y') = \frac{|Y'|}{|Y|}$, and when $E \subseteq X \times Y$, $\mu(E) = \frac{|E|}{|X|\cdot|Y|}$.

When $E \subseteq X \times Y$, for $x \in X$, we write $E_x = \{y \mid (x,y) \in E\}$ and for $y \in Y$ we write $E^y = \{x \mid (x,y) \in E\}$.

Definition 2.1. Let $E \subseteq X \times Y$. If $I \subseteq Y$, we say $\{E_x\}_{x\in X}$ *shatters* I if for each $J \subseteq I$, there is an $x \in X$ with $E_x \cap I = J$.

The *VC dimension* of a collection $\{E_x\}_{x\in X}$ is the supremum of $|I|$ for those $I \subseteq Y$ such that $\{E_x\}_{x\in X}$ shatters I. The *dual VC dimension* of $\{E_x\}_{x\in X}$ is the VC dimension of $\{E^y\}_{y\in Y}$.

When $E \subseteq X \times Y$, the VC dimension of E is the larger of the VC dimension of $\{E_x\}_{x\in X}$ and the dual VC dimension of $\{E_x\}_{x\in X}$.

Date: April 24, 2013.

Partially supported by NSF grant DMS-1340666.

An equivalent definition of the dual VC dimension is the supremum of $|I|$ where $I \subseteq X$ and for each $J \subseteq I$, there is a $y \in Y$ such that $y \in E_x$ iff $x \in J$. The definition of VC dimension in terms of the collection $\{E_x\}_{x\in X}$ is the standard one, but for us it will generally be more convenient to view VC dimension as a property of the set of pairs E (and we are only interested in the symmetric case where we look at the larger of the VC and dual VC dimensions).

Recall the following standard properties of VC dimension:

Lemma 2.2.
(1) *If $\{E_x\}_{x\in X}$ has VC dimension d, the dual VC dimension of $\{E_x\}_{x\in X}$ is less than 2^{d+1},*
(2) *If $X' \subseteq X$ and $Y' \subseteq Y$ then $E \cap (X' \times Y')$ has VC dimension no larger than the VC dimension of E,*
(3) [Shelah-Sauer [8, 9]] *If $\{E_x\}_{x\in X}$ has VC dimension d then for any set $I \subseteq Y$ with $|I| \geq d$,*
$$|\{E_x \cap I \mid x \in X\}| = |\{J \subseteq I \mid \exists x \in X\ J = E_x \cap I\}| \leq \left(\frac{e|I|}{d}\right)^d.$$

Theorem 2.3 (ϵ-net Theorem [3, 4])**.** *If $\{E_x\}_{x\in X}$ has VC dimension at most d, $d \geq 2$, then for sufficiently large $r \geq 2$ there is a set $\hat{Y} \subseteq Y$ such that $|\hat{Y}| \leq \mathrm{O}\big(dr \ln r\big)$ and for each $x \in X$ with $\mu(E_x) \geq 1/r$, $E_x \cap \hat{Y}$ is non-empty.*

Definition 2.4. We say $\hat{Y} \subseteq Y$ is an *ϵ-net for differences* if whenever $E_x \cap \hat{Y} = E_{x'} \cap \hat{Y}$, $\mu(E_x \triangle E_{x'}) < \epsilon$.

Lemma 2.5. *If $\{E_x\}_{x\in X}$ has VC dimension at most d, $d \geq 2$, then for each $r \geq 2$ there is an ϵ-net for differences $\hat{Y} \subseteq Y$ such that $|\hat{Y}| \leq \mathrm{O}\big(dr \ln r\big)$.*

Proof. For $x, x' \in X$, write $E_{x-x'} = E_x \setminus E_{x'}$. By [11, 1], the VC dimension of $\{E_{x-x'}\}_{x,x'\in X}$ is bounded by $10d$, so by the ϵ-net theorem, there is a set $\hat{Y} \subseteq Y$ such that $|\hat{Y}| \leq \mathrm{O}\big(dr \ln 2r\big) = \mathrm{O}\big(dr \ln r\big)$ so that whenever $\mu(E_{x-x'}) \geq 1/2r$, $E_{x-x'} \cap \hat{Y}$ is non-empty. In particular, if $E_x \cap \hat{Y} = E_{x'} \cap \hat{Y}$ then
$$\mu(E_x \triangle E_{x'}) = \mu(E_x \setminus E_{x'}) + \mu(E_{x'} \setminus E_x) \leq 1/2r + 1/2r = 1/r.$$
□

3. Regularity

In this section we fix a set $E \subseteq X \times Y$. We write $d(X', Y') = \frac{\mu(E \cap (X', Y'))}{\mu(X')\mu(Y')}$.

Throughout this section, a partition $\mathcal{P}$ is a pair $(\{X_i\}_{i \leq n}, \{Y_j\}_{j \leq m})$ where $\{X_i\}_{i\leq n}$ is a partition of X and $\{Y_j\}_{j\leq m}$ is a partition of Y. We set $|\mathcal{P}| = \max\{n, m\}$.

Definition 3.1. We say $\mathcal{P}' = (\{X'_i\}, \{Y'_j\})$ *refines* $\mathcal{P} = (\{X_i\}, \{Y_j\})$ if for each X_i and each $X'_{i'}$, either $X'_{i'} \subseteq X_i$ or $X'_{i'} \cap X_i = \emptyset$, and if for each Y_j and each $Y'_{j'}$, either $Y'_{j'} \subseteq Y_j$ or $Y'_{j'} \cap Y_j = \emptyset$.

Equivalently, $\mathcal{P}'$ refines $\mathcal{P}$ if for each X_i, there is some I' such that $\{X_{i'}\}_{i'\in I'}$ is a partition of X_i, and similarly for each Y_j. Clearly this relation is symmetric and transitive.

Definition 3.2. We define

$$\rho(\mathcal{P}) = \sum_{i\leq n, j\leq m} d^2(X_i, Y_j)\mu(X_i)\mu(Y_j).$$

We recall the following standard facts about ρ:

Lemma 3.3. *(1) $0 \leq \rho(\mathcal{P}) \leq 1$,*
(2) If $\mathcal{P}'$ refines $\mathcal{P}$ then $\rho(\mathcal{P}) \leq \rho(\mathcal{P}')$.

Definition 3.4. A pair (X_i, Y_j) is *ϵ-regular* if whenever $X' \subseteq X_i$, $Y' \subseteq Y_j$ with $\mu(X') \geq \epsilon\mu(X_i)$ and $\mu(Y') \geq \epsilon\mu(Y_j)$, we have

$$\left|d(X_i, Y_j) - d(X', Y')\right| < \epsilon.$$

If $\mathcal{P}$ is a partition, we write $\mathfrak{I}(\epsilon, \mathcal{P})$ for the set of (i, j) such that (X_i, Y_j) is *not* ϵ-regular.

We say $\mathcal{P}$ is ϵ-regular if $\sum_{(i,j)\in\mathfrak{I}(\epsilon,\mathcal{P})} \mu(X_i)\mu(Y_j) < \epsilon$.

Definition 3.5. Let $\hat{X} \subseteq X, \hat{Y} \subseteq Y$ be finite sets. The partition *induced by $\hat{X}, \hat{Y}$* takes, for each $I \subseteq \hat{Y}$, $X_I = \{x \mid E_x \cap \hat{Y} = I\}$ and for $J \subseteq \hat{X}$, $Y_J = \{y \mid E^y \cap \hat{X} = J\}$.

Lemma 3.6. *Suppose $\hat{X}, \hat{Y}$ are ϵ-nets for differences and let $\mathcal{P} = (\{X_i\}, \{Y_j\})$ be the partition induced by $\hat{X}, \hat{Y}$. Then whenever $x, x' \in X_i$, we have $\mu(E_x \triangle E_{x'}) < \epsilon$ and whenever $y, y' \in Y_j$, we have $\mu(E^y \triangle E^{y'}) < \epsilon$.*

Lemma 3.7. *Let $\hat{X} \subseteq \hat{X}' \subseteq X, \hat{Y} \subseteq \hat{Y}' \subseteq Y$ be given. Then the partition induced by $\hat{X}', \hat{Y}'$ refines the partition induced by $\hat{X}, \hat{Y}$.*

Lemma 3.8. *Suppose $\mathcal{P}$ is the partition induced by $\hat{X}, \hat{Y}$ and is not $1/r$-regular. Further, suppose E has VC dimension at most d. Then there are $\hat{X}' \supseteq \hat{X}$ and $\hat{Y}' \supseteq \hat{Y}$ such that the partion $\mathcal{P}'$ induced by $\hat{X}', \hat{Y}'$ satisfies:*

(1) $|\mathcal{P}'| \leq \mathrm{O}\big((|\mathcal{P}|dr^3 \ln r^3)^d\big)$,
(2) $\rho(\mathcal{P}') \geq \rho(\mathcal{P}) + 1/10^3 r^7$.

Proof. Let $\mathcal{P}$ be the partition $(\{X_i\}_{i\leq n}, \{Y_j\}_{j\leq m})$. For each i, define a measure μ_i on subsets of X_i by $\mu_i(X') = \frac{\mu(X')}{\mu(X_i)} = \frac{|X'|}{|X_i|}$. Similarly, define a measure μ^j on subsets of Y_j by $\mu^j(Y') = \frac{\mu(Y')}{\mu(Y_j)} = \frac{|Y'|}{|Y_j|}$. Finally, define a measure μ_i^j on subsets of $X_i \times Y_j$ by $\mu_i^j(S) = \frac{\mu(S)}{\mu(X_i)\mu(Y_j)}$.

For each $i \leq n$, let $\hat{X}'_i$ be a $1/10r^3$-net for differences in X_i with respect to the measure μ_i; take $\hat{X}' = \hat{X} \cup \bigcup_{i\leq n} \hat{X}'_i$. Similarly, for each $j \leq m$, let $\hat{Y}'_j$ be a $1/10r^3$-net for differences in Y_j with respect to the measure μ^j. By Lemma 2.5, each $\hat{X}'_i$ and $\hat{Y}'_j$ may be taken to have size at most $\mathrm{O}\big(dr^3 \ln r^3\big)$.

This means $|\hat{X}'|, |\hat{Y}'| \leq \mathrm{O}\big(|\mathcal{P}|dr^3 \ln r^3\big)$. Let $\mathcal{P}'$ be the partition induced by $\hat{X}', \hat{Y}'$. By Shelah-Sauer, $|\mathcal{P}'| \leq \mathrm{O}\big(\big(|\mathcal{P}|dr^3 \ln r^3\big)^d\big)$.

It remains to show that

$$\rho(\mathcal{P}') \geq \rho(\mathcal{P}) + 1/10^3 r^7.$$

For each i, j, $\mathcal{P}'$ induces a partition $\mathcal{P}'_{i,j}$ of X_i, Y_j, and we have

$$\rho(\mathcal{P}') = \sum_{i \leq n, j \leq m} \rho_{i,j}(\mathcal{P}'_{i,j})\mu(X_i)\mu(Y_j)$$

where

$$\rho_{i,j}(\mathcal{P}'_{i,j}) = \sum_{X'_{i'} \subseteq X_i, Y'_{j'} \subseteq Y_j} d^2(X'_{i'}, Y'_{j'})\mu_i(X'_{i'})\mu^j(Y'_{j'}).$$

So it suffices to show that whenever $(i,j) \in \mathfrak{I}(1/r, \mathcal{P})$,

$$\rho_{i,j}(\mathcal{P}'_{i,j}) \geq d^2(X_i, Y_j) + 1/10^3 r^6.$$

So consider some $(i,j) \in \mathfrak{I}(1/r, \mathcal{P})$, and let $\mathcal{P}'_{i,j} = (\{X'_{i'}\}_{i' \leq n'}, \{Y'_{j'}\}_{j' \leq m'})$. Let $X' \subseteq X_i, Y' \subseteq Y_j$ witness the failure of $1/r$-regularity. That is, $\mu_i(X') \geq 1/r$, $\mu^j(Y') \geq 1/r$, and

$$\big|d(X', Y') - d(X_i, Y_j)\big| \geq 1/r.$$

Assume $d(X', Y') \geq d(X_i, Y_j) + 1/r$ (the case where $d(X', Y') \leq d(X_i, Y_i) - 1/r$ is symmetric). Let $\tilde{X}'$ consist of those $x \in X_i$ such that $\frac{\mu^j(E_x \cap Y')}{\mu^j(Y')} \geq d(X_i, Y_j) + 1/2r$; clearly $\frac{\mu_i(\tilde{X}')}{\mu_i(X')} \geq 1/2r$, and so $\mu_i(\tilde{X}') \geq 1/2r^2$. Set

$$\tilde{X} = \bigcup \{X'_{i'} \mid X'_{i'} \cap \tilde{X}' \neq \emptyset\}.$$

Whenever $x \in \tilde{X}$, we have $x \in X'_{i'}$ and some $x' \in \tilde{X}'$ so that $\mu_j((E_x \triangle E_{x'}) \cap Y_j) < 1/10r^3$. In particular, since $\frac{\mu^j(E_{x'} \cap Y')}{\mu^j(Y')} \geq d(X_i, Y_j) + 1/2r$ and $\mu^j(Y') \geq 1/r$, $\frac{\mu^j(E_x \cap Y')}{\mu^j(Y')} \geq d(X_i, Y_j) + 2/5r$. Note that $\tilde{X}' \supseteq \tilde{X}$, and so $\mu_i(\tilde{X}) \geq 1/2r^2$.

Now let $\tilde{Y}'$ consist of those $y \in Y_j$ such that $\frac{\mu_i(E^y \cap \tilde{X})}{\mu_i(\tilde{X})} \geq d(X_i, Y_j) + 1/5r$. Clearly $\mu^j(\tilde{Y}') \geq 1/5r^2$. Set

$$\tilde{Y} = \bigcup \{Y'_{j'} \mid Y'_{j'} \cap \tilde{Y}' \neq \emptyset\}.$$

Again, whenever $y \in \tilde{Y}$ we have a y' with $\mu^i((E^y \triangle E^{y'}) \cap X_i) < 1/10r^3$ and $\frac{\mu_i(E^y \cap \tilde{X})}{\mu_i(\tilde{X})} \geq d(X_i, Y_j) + 2/5r$, and therefore $\frac{\mu_i(E^y \cap \tilde{X})}{\mu_i(\tilde{X})} \geq d(X_i, Y_j) + 1/10r$. It follows that $d(\tilde{X}, \tilde{Y}) \geq d(X_i, Y_j) + 1/10r$.

Consider the partition $\mathcal{P}^*_{i,j} = (\{\tilde{X}, X_i \setminus \tilde{X}\}, \{\tilde{Y}, Y_j \setminus \tilde{Y}\})$. $\mathcal{P}'_{i,j}$ refines $\mathcal{P}^*_{i,j}$ (since $\tilde{X}, \tilde{Y}$ were defined to be unions of components from $\mathcal{P}'_{i,j}$), so it suffices to show that

$$\rho_{i,j}(\mathcal{P}^*_{i,j}) \geq d^2(X_i, Y_j) + 1/10^3 r^6.$$

Since $\mu(\tilde{X})\mu(\tilde{Y}) \geq 1/10r^4$ and $d(\tilde{X}, \tilde{Y}) \geq d(X_i, Y_j) + 1/10r$, this follows from standard calculations.

$\square$

Theorem 3.9. *If E has VC dimension at most d, there is a $1/r$-regular partition $\mathcal{P}$ with $|\mathcal{P}| \leq \mathrm{O}\big((dr^3 \ln r^3)^{d^{2\cdot 10^3 r^7}}\big)$.*

Proof. Let $\mathcal{P}_0$ be the trivial partition $(\{X\}, \{Y\})$. Given $\mathcal{P}_i$ not $1/r$-regular, the previous lemma tells us there is a $\mathcal{P}_{i+1}$ refining $\mathcal{P}_i$ with $|\mathcal{P}_{i+1}| \leq \mathrm{O}\big((|\mathcal{P}_i| dr^3 \ln r^3)^d\big)$ and $\rho(\mathcal{P}_{i+1}) \geq \rho(\mathcal{P}_i) + 1/10^3 r^7$. Since $0 \leq \mathcal{P}_0$ and $\mathcal{P}_n \leq 1$, we must have $n \leq 10^3 r^7$, and therefore there is an $n \leq 10^3 r^7$ with $\mathcal{P}_n$ $1/r$-regular.

It is easy to check inductively that

$$|\mathcal{P}_i| \leq \mathrm{O}\big((dr^3 \ln r^3)^{d^{2i}}\big).$$

$\square$

References

[1] R. M. Dudley. A course on empirical processes. In *École d'été de probabilités de Saint-Flour, XII—1982*, volume 1097 of *Lecture Notes in Math.*, pages 1–142. Springer, Berlin, 1984.

[2] W. T. Gowers. Lower bounds of tower type for Szemerédi's uniformity lemma. *Geom. Funct. Anal.*, 7(2):322–337, 1997.

[3] David Haussler and Emo Welzl. ϵ-nets and simplex range queries. *Discrete Comput. Geom.*, 2(2):127–151, 1987.

[4] János Komlós, János Pach, and Gerhard Woeginger. Almost tight bounds for ϵ-nets. *Discrete Comput. Geom.*, 7(2):163–173, 1992.

[5] Michael C. Laskowski. Vapnik-Chervonenkis classes of definable sets. *J. London Math. Soc. (2)*, 45(2):377–384, 1992.

[6] László Lovász and Balázs Szegedy. Regularity partitions and the topology of graphons. In *An irregular mind*, volume 21 of *Bolyai Soc. Math. Stud.*, pages 415–446. János Bolyai Math. Soc., Budapest, 2010.

[7] M. Malliaris and S. Shelah. Regularity lemmas for stable graphs. *ArXiv e-prints*, February 2011.

[8] N. Sauer. On the density of families of sets. *J. Combinatorial Theory Ser. A*, 13:145–147, 1972.

[9] Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41:247–261, 1972.

[10] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.

[11] Aad van der Vaart and Jon A. Wellner. A note on bounds for VC dimensions. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. Collect.*, pages 103–107. Inst. Math. Statist., Beachwood, OH, 2009.

Department of Mathematics, University of Pennsylvania, 209 South 33rd Street, Philadelphia, PA 19104-6395, USA

E-mail address: `htowsner@math.upenn.edu`

URL: `www.math.upenn.edu/~htowsner`